

S'il suffisait d'étiqueter un bon de commande avec une balise «commande» pour que l'ordinateur puisse le traiter, nous n'aurions pas besoin de XML. Nous n'aurions même pas besoin de programmeurs, les ordinateurs se débrouilleraient tout seuls.

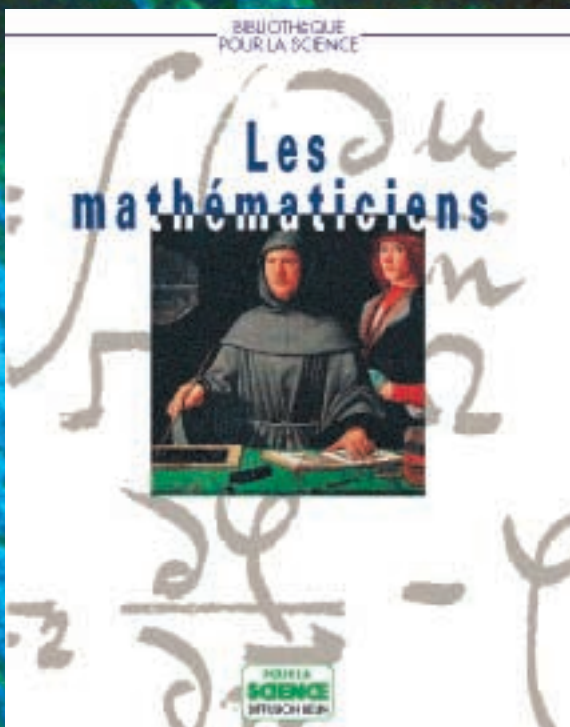
Le langage XML n'est pas magique, mais il est efficace ; il définit des règles de base qui éliminent une couche de programmation correspondant à des détails, de sorte que des personnes partageant les mêmes intérêts peuvent se concentrer sur la véritable difficulté : s'entendre sur la façon dont ils souhaitent représenter l'information qu'ils échangent couramment. Ce problème n'est pas simple, mais il n'est pas nouveau.

Ces harmonisations sont incontournables, car la prolifération de systèmes informatiques incompatibles a engendré des retards, des coûts et de la confusion dans la plupart des activités humaines. Les utilisateurs de l'informatique veulent échanger des idées ou faire des affaires sans utiliser tous les mêmes ordinateurs ; des langages d'échange spécifiques à chaque activité y contribuent grandement. De fait, l'avalanche d'acronymes se terminant par ML témoigne de la créativité qui s'exprime grâce à XML en sciences ou dans le monde des affaires (voir l'encadré de la page 52).

Les concepteurs d'un nouveau langage doivent s'accorder sur trois points : quelles balises seront acceptées, comment les éléments balisés pourront être emboîtés et comment ils seront traités. Les deux premières spécifications, le vocabulaire du langage et sa structure, sont codifiées dans une définition de type de document (ou DTD). Le standard XML n'impose pas de telles définitions, mais la plupart des nouveaux langages en comporteront probablement, parce qu'elles simplifient le travail des programmeurs. Ces derniers doivent en effet écrire des programmes qui comprennent le balisage et s'en servent intelligemment.

Les programmeurs auront également besoin d'un ensemble de directives qui décrivent en langage intelligible ce que signifient toutes les balises. Le langage HTML, par exemple, est défini par une DTD et par quelques centaines de pages de descriptions que consultent les programmeurs quand ils écrivent des programmes de navigation ou d'autres programmes pour le réseau Internet.

BIBLIOTHÈQUE POUR LA SCIENCE



LES MATHÉMATICIENS

Pour comprendre le processus de la création mathématique, il faut saisir ce qui fait l'originalité d'un mathématicien. Cet ouvrage réunit les portraits de quinze grands mathématiciens. Leur rapport avec leur époque, le regard qu'ils portent sur leur activité sont plus qu'anecdotiques : ils révèlent la genèse des idées nouvelles et mettent en lumière le caractère audacieux, dynamique et imaginaire de l'invention mathématique.

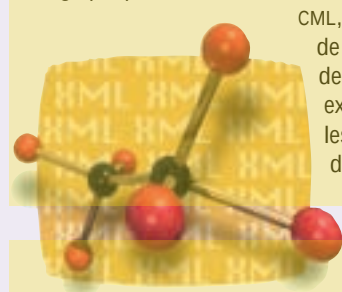
Voir bon de commande p. 98

POUR LA
SCIENCE
DIFFUSION BELIN

Les nouveaux langages des sciences

Grâce à XML, les scientifiques échangent plus facilement des théories, des calculs et des résultats expérimentaux. Depuis longtemps, les mathématiciens étaient frustrés par l'inaptitude des navigateurs à afficher les expressions mathématiques autrement que sous forme d'images. Le langage MathML, opérationnel depuis le printemps 1999, leur permet de couper des équations dans les pages Internet et de les coller directement dans leurs logiciels de calcul formel, en vue de les utiliser pour des calculs ou pour des graphiques.

Les chimistes vont plus loin encore : ils ont mis au point de nouveaux navigateurs pour leur langage CML (*Chemical Markup Language*), fondé sur XML, qui donne des représentations graphiques de la structure moléculaire des composés décrits dans les pages Internet.



CML, de même que le langage de l'astronomie, en cours de mise au point, aide les chercheurs à explorer rapidement une multitude de citations de revues pour en extraire les articles correspondant à leur sujet d'étude : les astronomes, par exemple, peuvent rentrer les coordonnées célestes d'une galaxie, afin de récupérer une liste d'images, d'articles et de données instrumentales sur l'objet en question.

XML sera utile pour la réalisation d'expériences aussi bien que pour l'analyse de leurs résultats. En 1998, les ingénieurs de la NASA ont entrepris de mettre au point l'AIML (*Astronomical Instrument ML*, soit « langage ML pour les instruments astronomiques »), qui commande le télescope infrarouge SOFIA lors de ses vols stratosphériques à bord d'un Boeing 747. À terme, AIML permettra aux astronomes du monde entier de commander des télescopes et, peut-être même, des satellites, par l'intermédiaire d'un simple logiciel de navigation sur le réseau Internet.

Pendant ce temps, sur Terre, les généticiens utiliseront bientôt BSML (*Biosequence ML*), afin d'échanger et de manipuler les quantités énormes d'information qui résultent des innombrables projets de séquençage du génome. Un navigateur BSML, mis au point et distribué gratuitement par la Société *Visual Genomics*, permet aux chercheurs de fouiller de gigantesques bases de données de code génétique et d'afficher les fragments trouvés sous forme de cartes et de tableaux, bien plus explicites que les suites de lettres habituelles.

Question de style

Les utilisateurs, eux, sont préoccupés de ce que font les programmes, et non de leur description. Ils veulent que les programmes affichent de manière intelligible l'information codée en XML, mais les balises XML ne donnent pas d'indications sur la présentation de l'information sur l'écran ou sur le papier.

Cette caractéristique est d'ailleurs un avantage pour les éditeurs, qui veulent souvent réutiliser le contenu d'un document, c'est-à-dire décliner une publication sous une multitude de formes, imprimées ou électroniques. En balisant le contenu d'un texte pour en décrire son sens, XML autorise cette publication multiple, indépendamment du type de support. Les éditeurs peuvent ensuite appliquer des règles organisées en feuilles de style pour mettre en forme automatiquement le passage sous diverses formes. Le standard pour les feuilles de style XML, en cours de mise au point, est XSL (pour *eXtensible Stylesheet Language*).

Les versions les plus récentes de certains navigateurs Internet peuvent lire un document XML, récupérer la feuille de style appropriée, et l'utiliser pour trier et formater l'information sur l'écran. L'utilisateur ne verra pas la différence entre des informations codées par HTML ou par XML, mais l'accès aux sites XML sera bien plus rapide.

Les malvoyants gagneront beaucoup de cette approche de l'édition, car les feuilles de style leur permettront de présenter un document XML en braille ou sous forme parlée. De même, les pages Internet pourront être lues à voix haute par les ordinateurs de bord des voitures (voir la figure 2).

Un traducteur universel

Bien que le réseau Internet ait été créé par le milieu universitaire et scientifique, c'est l'industrie et le commerce qui, attirés par les potentialités du système, ont suscité son expansion fulgurante. Le commerce électronique s'est imposé, mais les transactions entre entreprises se développent au

moins aussi rapidement. Les industries aimeraient utiliser le réseau pour régler leur production sur les commandes qu'elles enregistrent, mais cet ajustement, qui limiterait les stocks, impose une automatisation des chaînes de production et des communications avec les fournisseurs de matières premières ou de pièces détachées. Or, les interactions de systèmes informatiques sont gênées par l'hétérogénéité des traitements.

Depuis des siècles, les affaires se font par l'échange de documents standardisés : bons de commande, factures, reçus. Ces documents assurent les interactions des parties concernées. Chaque document contient exactement ce que son destinataire a besoin de savoir, et rien de plus. De même, l'échange de documents reste le fondement des transactions commerciales en ligne, mais le langage HTML n'a pas été conçu pour ces échanges. En revanche, XML a été conçu pour l'échange de documents, et nous ne doutons pas que le commerce électronique mondial reposera largement sur la circulation de transactions matérialisées par des millions de documents XML échangés par l'Internet.

Ainsi, le réseau renforcé par XML sera plus rapide, plus convivial et mieux adapté aux affaires. En contrepartie, les concepteurs des sites auront plus de travail. Des bataillons de programmeurs seront nécessaires pour tirer un plein parti des nouveaux langages XML. Bien que les jours des autodidactes du réseau ne soient pas encore comptés, l'espèce est en danger. Les concepteurs de demain devront être versés non seulement dans la production de textes et de graphiques, mais encore dans la construction de systèmes combinant à plusieurs niveaux des arbres de données, des structures de liens hypertexte, des métadonnées et des feuilles de style, qui, ensemble, constituent l'infrastructure robuste du réseau de deuxième génération.

Jon BOSAK, qui travaille pour la Société *Sun Microsystems*, est président du groupe de coordination XML du Consortium W3C. Tim BRAY, fondateur de la Société *Textuality*, est coéditeur de la norme XML 1.0.

Alain MICHARD, *XML : langage et applications*, Eyrolles, 1998.

Des liens pour cet article sont disponibles sur notre site : <http://www.pourlascience.com>